Baseline Levels of Service and Cost Estimation: Discussion

Vanessa Griffin

SEEDS Deputy Formulation Manager

Vanessa.l.griffin.1@gsfc.nasa.gov

301.614.5365

Agenda

- Baseline Levels of Service
- LOS – Cost Estimation Study Goals
- Approach
- Concepts
- Milestones
- Workshop Goals for LOS – Cost Estimation Study
- Plan for Breakout Sessions
  - Topics for Provider Services Session
  - Topics for Science Discipline Session
  - Topics for LOS / Cost Study Team Breakouts
- Background: Functional / Cost Areas
- Background: Strawman Levels of Service

Baseline Levels of Service

- This workshop will be primarily focused on establishing a baseline set of levels of service that users will need from the SEEDS data service providers (and that providers need from each other).
  - The baseline will be fed into guidelines and recommendations to be provided to ESE by the SEEDS office.
  - Minimum level of service for the user community expected; higher levels of service may be recommended or desirable.
- The ability to meet these recommended levels of service could be a factor in the future selection of service providers as well as the basis for SEEDS performance measurement.
- The baseline levels of service could vary depending on type of service provided as well as potential multiple-discipline use of the data.
- The baseline is needed by the Formulation team to plan for effective data systems as well as estimate the costs for ESE science data management over the next ten years.
- To support cost estimation, levels of service are associated with a set of draft requirements that are intended to encompass all data service provider functions and areas of cost.

3

Levels of Service and Cost Estimation Study Goals

Goal: Develop data service provider levels of service:

- Levels of Service (LOS) must be driven by science community – program science, project science, researchers.
- Use a draft set of requirements associated with LOS to support cost estimation; not necessarily requirements to be levied on data service providers by NASA, e.g. they could be similar to requirements owned by the provider.
- LOS and associated requirements will not embody or drive an architecture.

Goal: Develop a cost estimation model that estimates life cycle costs for data service providers:

- Model to be driven by levels of service and requirements;
- Model to support estimating individual data service provider costs and architecture trades.
- To facilitate cost estimation and support architecture trades, develop a set of data service provider types (architecture components) map LOS-requirements set to these.

Approach

1. Data service provider cost estimation must be grounded on an understanding of baseline levels of service and requirements.

- Define a set of 'functional areas' that span all significant areas of cost.
- Develop a general Data Services Provider Reference Model that relates levels of service / requirements and model parameters (metrics) for the functional areas.
- Develop a strawman set of general levels of service / requirements for community review – refine levels of service to a baseline established with the community at this workshop.
- Develop a working set of data service provider types – architecture components – map levels of service and associated requirements to these.
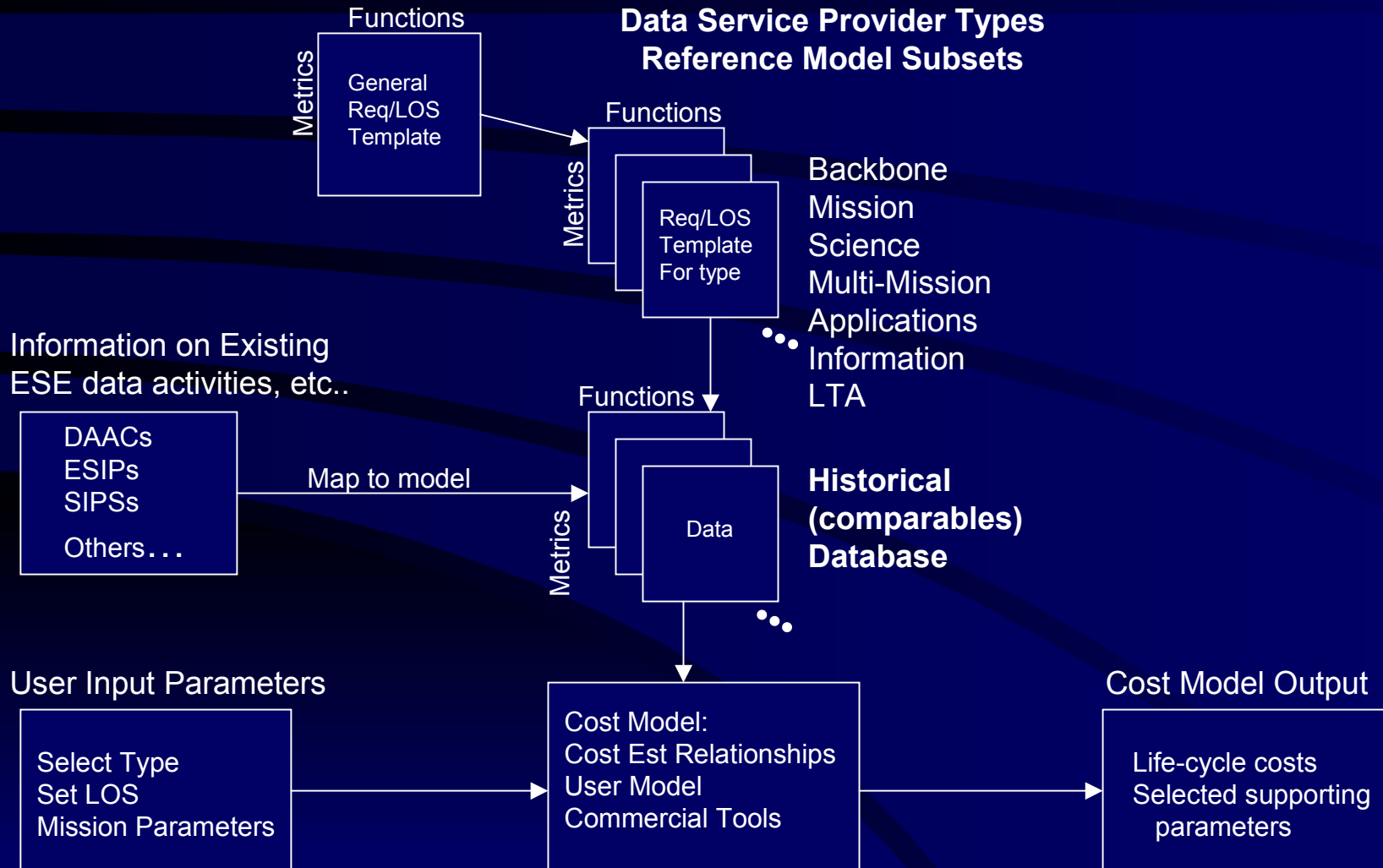
2. Use Cost Estimating by Analogy as basic methodology, augment with commercial cost estimating tools.

- Build database of 'comparables' – existing ESE and other data activities, bin by data service provider type and functionality.
- Develop cost estimating relationships through analysis of the comparables data base.
- Produce test/prototype data service provider life cycle cost estimates.
- Test against independent cases, expand database, revise and improve model. Provide prototype for evaluation, followed by improved versions.
- Provide data service provider cost estimates, single and architectures.
- Provide a portable capability to enable users to generate their own cost estimates.

Concepts: Requirements/LOS, Data Service Provider Reference Model, Cost Estimation

General Data Service Provider Reference Model

Functions

Metrics

General Req/LOS Template

Data Service Provider Types Reference Model Subsets

Functions

Metrics

Req/LOS Template For type

Backbone
Mission
Science
Multi-Mission
Applications
Information
LTA

Information on Existing ESE data activities, etc..

DAACs
ESIPs
SIPSs

Others…

Map to model

Functions

Metrics

Data

Historical (comparables) Database

User Input Parameters

Select Type
Set LOS
Mission Parameters

Cost Model:
Cost Est Relationships
User Model
Commercial Tools

Cost Model Output

Life-cycle costs
Selected supporting parameters

Study Milestones

• October 1, 2001 - Start

• January 16, 2002 - Strawman Requirements / Levels of Service

• January 2002 - Began building comparables database

• February 5-7, 2002 – Community Workshop

• March 31, 2002 - Revised Requirements / Levels of Service

• June 1, 2002 - Prototype Cost Estimation Model

• June 2002 - Community Workshop – Evaluate Prototype Cost Model

Continue database building, model evaluation and refinement…

Workshop Goals for
LOS & Cost Estimation Study

This workshop will be primarily focused on establishing a baseline set of levels of service that users will need from the SEEDS data service providers.

We need your review, evaluation, critique, recommended changes:

- 1. Strawman general data service provider requirements / levels of service (Workshop Paper 1);
- 2. Working list of data service provider types (Workshop Paper 2);
- 3. Mapping of general requirements / LOS set to data service provider types (Workshop Paper 3).

Your feedback:

White papers submitted, comment sheets filled out, will be used and responded to by email. Notes will be taken in breakouts.

Your feedback will be incorporated into the new baseline levels of service - you will see it happen on the website, you will see results at future workshops.

Plan for Breakout Sessions

Three sets of breakouts:

Objective is to make sure that levels of service and functionality meet:
• Users' research needs;
• Data Service Providers' needs from each other, so that providers can meet users' needs.

Wednesday Morning: Services needed from/by providers
• Users from all science disciplines and data service providers look at individual data service provider types.

Wednesday Afternoon: Services needed by science disciplines
• Users from individual science disciplines and data service providers look at all data service provider types.

Wednesday Afternoon / Thursday AM: LOS and Cost Estimation Study
• In parallel with other study team breakouts.
• Anyone interested in more detail on study approach, methodology.

Topics for Provider Services Breakout Sessions

- For a given type of data service provider from the perspective of all users:

- 1. Is our approach (minimum, recommended, desired) to setting levels of service appropriate from your perspective? Are there better approaches you could suggest?

- 2. Are the minimum levels of service for this provider type adequate to meet your research needs?

- 3. Are there changes you would suggest in the classification of this provider type's levels of service as minimum, recommended, desirable?

- 4. Are there additional or better levels of service for this provider type, or any that seem unreasonable to you as a user?

- 5. Is the mapping of requirements and levels of service to this provider type reasonable? Does the mapping reflect accurately the role of this provider type?

Topics for Science Discipline Breakout Sessions

For all types of data service providers from the perspective of a science discipline:

1. Is our approach (minimum, recommended, desired) to setting levels of service appropriate from your perspective? Are there better approaches you could suggest?

2. How well does this set of data service provider types meet your research needs? Is there a need for additional types, or for additional functions in one or more types, to meet the needs of your science discipline?

3. Taken together, are the minimum levels of service for the set of provider types adequate to meet your research needs?

4. Are there changes you would suggest in the classification of any of the provider type's levels of service as minimum, recommended, desirable?

5. Are there additional or better levels of service for any of the data provider types, or any that seem unreasonable to you as a user?

6. Is the mapping of requirements and levels of service to the set of data service provider types reasonable? Does the mapping reflect accurately the role of the data service providers? Are there any gaps that you see?

Topics for LOS/Cost Team Breakout Sessions

1. Is there an area of cost or cost factor that is not addressed, or not addressed specifically enough, in the requirements / levels of service?

2. From a functional point of view, are there areas of functionality that are not addressed or not addressed specifically enough?

3. Are the seven ESE data service provider types representative of the different types of institutional components / data centers / etc. that will be implemented in the next ten years? If not, can you suggest others?

4. Is the mapping of requirements and levels of service to data provider types reasonable? Does the mapping reflect accurately the different roles of the different data service providers?

5. For data service providers, does our two step approach to building the 'comparables' database seem reasonable (we research documentation, websites etc. first, follow up with you for information we can't find)?

Functional / Cost Areas 1

- Ingest—the process of receiving, reading, quality checking, cataloging, of incoming data to the point of insertion into the archive. Ingest can include format conversion, metadata extraction, etc.

- Processing—the generation and quality checking of new derived data products from data or products that have been ingested, or previously generated. Processing includes process control (production planning, scheduling, monitoring, etc.) as well as product generation per se.

- Documentation---the development (or upgrading of received) data and product documentation to meet adopted standards, including catalog information (metadata), user guides, etc.

- Archive—the insertion of data into archive storage, and management, handling and preservation of data, metadata, and documentation within the archive.

- Distribution—providing access to catalog information and a search &order capability, receiving user requests, fetching the requested items from the archive, performing any subsetting, reformatting / format conversion, or packaging, and providing the end product to the user by network or media.

- User Support—Support provided in direct contact with users by user support staff, including responding to queries, taking of orders, etc.

- Instrument / Mission Operations - monitoring instrument (spacecraft) performance, generating instrument (spacecraft) commands, and event scheduling.

14

Functional / Cost Areas 2

- Sustaining Engineering - Maintenance / enhancement of custom software.

- Engineering Support - Systems engineering, test engineering, configuration management, system administration, database administration, network engineering, etc. Directed toward the internal operation of the data service provider.

- Technical Coordination – Participation in system level processes, including coordination on standards and interfaces, common metrics, etc. Directed outward, supporting the data service provider as one of a system of cooperating centers.

- Implementation---Includes design, development of, and making operational, required data and information system required capabilities. Recurs as systems are expanded or replaced.

- Management - Includes management and administration at the data service provider level ("front office") and direct management of functional areas, and internal support such as logistics, supplies, facilities, security management, property inventory and management, facility management.

- Facility / Infrastructure - Non-staff cost factors such as supplies, facility lease and utility costs, hardware maintenance, COTS licenses, etc.

Strawman Levels of Service - Ingest

- Ingest
- Ingest satellite data streams, ancillary data products, processed products generated by other data service providers, etc.
- Levels of Service:
- 1) operational (time-critical) ingest with immediate verification of data integrity and quality;
- 2) routine ingest and verification of data quality and integrity without tight time constraints;
- 3) ad hoc or intermittent ingest on a non-operational basis with verification of data quality and integrity;
- 4) ad hoc or intermittent ingest on a non-operational basis.

Strawman Levels of Service – Processing

Processing

Standard product generation, on a routine schedule or on-demand.

Levels of Service:

1) standard products generated within 2 days of ingest/availability of required inputs;
2) within 7 days of ingest/availability of required inputs;
3) within 30 days of ingest/availability of required inputs.

Ad hoc, non-operational product generation.

Levels of Service:

1) specific targets for processing adopted on a case by case basis;
2) general goals for processing;
3) no goals, purely ad hoc processing.

Reprocessing of standard products in response to reprocessing requests.

Levels of Service:

1) the capacity for reprocessing at 9 times the original processing rate;
2) at 6 times the original processing rate;
3) at 3 times the original processing rate.

Strawman Levels of Service - Documentation

Documentation

Provide standard compliant catalog information (metadata, including browse) and documentation describing all data and information.

Levels of Service:

1) data and product holdings documented to the standard for long term archiving;

2) documentation ensured to be sufficient for current use;

3) documentation only as received from product provider.

Update documentation of data and products with user comments.

Levels of Service:

1) data and products routinely updated with user comments;

2) data and products occasionally updated with user comments;

3) data and products rarely updated with user products.

Strawman Levels of Service - Archive

Archive / working storage capacity of [number] TB.

Levels of Service:

1) archive capacity is cumulative sum of all data ingested plus all products generated;

2) archive capacity is limited to a specified threshold.

Quality screening on data entering and/or exiting the archive.

Levels of service:

1) exit and entry screening;

2) entry screening.

Preservation of archived data – screening to detect media problems.

Levels of service:

1) 10% per year random screening;

2) 5% per year random screening;

3) 1% per year random screening.

Robust archive media.

Levels of Service:

1) archive media compliant with NARA standards;

2) archive media consistent with commercial practice.

Strawman Levels of Service - Archive, Continued

Archive / working storage backup

Levels of service:

1) full off-site backup, with regular sampling to verify integrity;

2) partial, [Backup Fraction - % of archive backed up], off-site backup, with sampling;

3) partial, [Backup Fraction - % of archive backed up], on-site backup with sampling.

Backup facility for its data and information holdings.

Levels of Service:

1) an environmentally controlled and physically secure off-site backup archive facility;

2) an on-site but separate environmentally controlled and physically secure off-site backup facility;

3) a backup capability within the data service provider's primary data system(s).

Periodic migration of archive to new archive media / technology.

Levels of Service:

1) planned migration;

2) no planned migration, but ad hoc migration as need is seen to arise.

Strawman Levels of Service - Distribution

Access to all metadata and data and information holdings, including all standard science products.

Levels of Service:

1) public access to all users;

2) access to the science community;

3) access to a limited team of scientists.

Search and order capability to [all users (including the general public) consistent with SEEDS standards and practices; to a limited set of science team members].

Levels of Service:

1) allow search for instances of product types by geophysical parameter, time, and space across multiple product types;

2) allow search for instances of multiple product types by time and space;

3) allow search for instances of single product type by time and space.

Strawman Levels of Service – Distribution, Continued

Perform timely distribution of data and products to users by network.

Levels of service:

1) availability of a product for network delivery within ten seconds;

2) availability of a product for network delivery within ten minutes;

3) availability of a product for network delivery within twenty four hours.

Perform timely distribution of data and products to users on SEEDS standard media types.

Levels of Service:

1) shipping of media product within three days of receipt of request;

2) shipping of media product within one week of receipt of request,

3) shipping of media product within one month of receipt of request.

Distribution services such as subsetting, reformatting, and packaging.

Levels of Service:

1) supporting services available for most archived data and products;

2) supporting services available for less than half of archived data and products;

3) supporting services available for a few selected data and products only.

Strawman Levels of Service - Other

User Support

The data service provider shall provide a trained user support staff.

Levels of service:

1) one user support staff member per 100 active users;

2) one user support staff member per 500 active users;

3) one user support staff member per 1,000 active users.

Sustaining Engineering

Maintain and enhance developed / reused custom software.

Levels of Service:

1) no or very infrequent interruptions of data service provider operations;

2) occasional interruptions in data service provider operations;

3) as needed, with interruptions in data service provider operations a secondary concern.

Strawman Levels of Service – Other Continued

Engineering Support

System administration, network administration, database administration, coordination of hardware maintenance by vendors, and other technical functions.

Levels of Service:

1) no or very infrequent interruptions of data service provider operations;

2) occasional interruptions in data service provider operations;

3) as needed, with interruptions in data service provider operations a secondary concern.

Systems engineering, test engineering, configuration management, COTS procurement, installation of COTS upgrades, network/communications engineering and other engineering.

Levels of Service:

1) no or very infrequent interruptions of data service provider operations;

2) occasional interruptions in data service provider operations;

3) as needed, with interruptions in data service provider operations a secondary concern.